

## “囚徒困境”的前世今生（下）

郝继仁 北京大学

### 一、引言

在《“囚徒困境”的前世今生》上篇和中篇中，我们介绍了博弈论的起源和发展，并对促进合作演化的三种重要机制——直接互惠、间接互惠以及网络互惠做了介绍。在这些机制的研究中常用的博弈范式为囚徒困境，其收益矩阵一般是如下所示的捐赠者博弈形式：

$$\begin{array}{c} C \\ D \end{array} \begin{array}{cc} C & D \\ \left( \begin{array}{cc} b-c & -c \\ b & 0 \end{array} \right) \end{array}$$

在该博弈中，合作者（C）付出代价 $c$ 给对手提供好处 $b$  ( $b > c > 0$ )，背叛者（D）不付出代价也不提供好处。参数 $b > c > 0$ 的设定使得单方面的背叛有利可图 ( $b > b - c, 0 > -c$ )，然而相互背叛导致博弈双方的收益均比相互合作低 ( $0 < b - c$ )，因此囚徒困境刻

画了个体利益与集体利益之间的冲突。

在本篇中，我们将着重介绍关于网络互惠研究的后续进展，包括交互网络与学习网络、多个体博弈、基于自我评估的决策规则以及动态网络。在详细介绍这些内容之前，我们先回顾一下《“囚徒困境”的前世今生》中篇中介绍过的网络互惠的内涵。网络互惠指的是合作者能在网络上与相邻的合作者形成团簇，互惠互利，以此来抵御背叛者的入侵（见图1）。该机制区别于直接互惠和间接互惠的关键点在于其所刻画的问题情境中，个体间的交互不再是随机进行的。具体来说，只有在网络上互为邻居的个体之间才能进行交互。正是在这种受限制的局部交互下，合作者之间才能形成紧密的团簇，从而在与背叛者的竞争中存活下来。

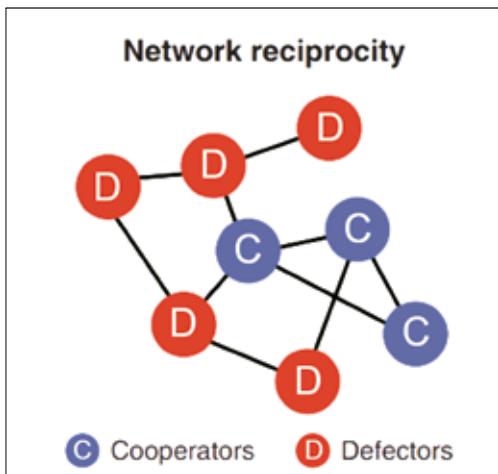


图1 网络互惠示意图。源自Nowak M A. Five rules for the evolution of cooperation [J]. Science, 2006, 314(5805): 1560-1563.

## 二、交互网络与学习网络

在Nowak和May等人1992年的研究<sup>[1]</sup>中，一个重要的假设是个体的交互对象（即个体进行博弈的对象）与学习对象（即个体进行策略更新时参考的对象）是完全相同的。这样的假设可能是与现实不符的，比如一个人日常的交互对象可能是他的同事、周围的街坊邻居，但他的学习对象可

能来自虚拟的社交网络。这说明交互网络和学习网络不一定是相同的，它们之间可能存在差异。

基于此，Ohtsuki 等人在 2007 年进行了建模，提出了交互网络和学习网络（也称取代网络）的概念：每个个体的交互对象是它交互网络上的邻居，而它的学习对象则是它学习网络上的邻居<sup>[2]</sup>。在他们的模型中，个体在进行策略更新时采用的是灭生规则（death-birth rule），即每一时间步一个随机的个体会死亡，然后它周围的所有邻居以正比于它们适应度（或者收益）的概率竞争这个空位，竞争成功的个体会产生一个相同的子代并占据这个空位。这个更新规则也可以理解为一种模仿学习，即一个随机的个体会放弃它自己当前的策略，并以正比于它邻居的收益的概率来学习它邻居的策略。此外，Ohtsuki 等人研究的网络为随机规则网络（random regular graphs），并且交互网络和学习网络拥有同样的节点。当交互网络的度为  $h$ ，学习网络的度为  $g$ ，并且假设它们相互重叠的边所形成的网络为一个度为  $l$  的随机规则网络时（见图 2），合作者被自然选择所青睐的条件是收益代价比  $b/c > hg/l$ 。这个研究结果表明，交互网络和学习网络的重合度越高（即  $l$  越大），合作越容易被促进。

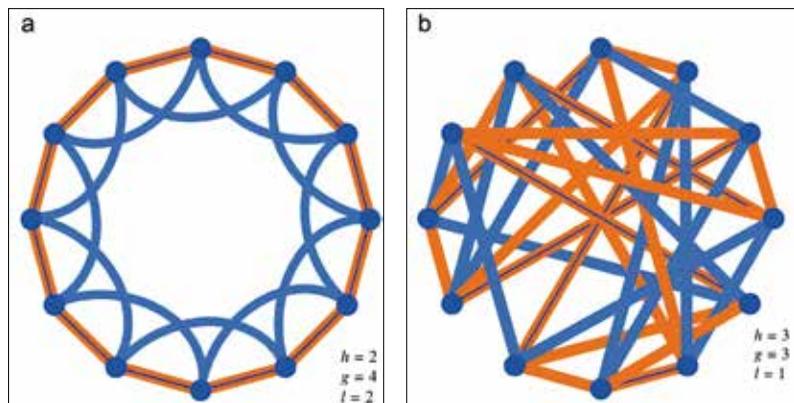


图2 交互网络（红色）和学习网络（蓝色）示意图。

源自Ohtsuki H, Pacheco J M, Nowak M A. Evolutionary graph theory: Breaking the symmetry between interaction and replacement [J]. Journal of Theoretical Biology, 2007, 246(4): 681-694.

### 三、多个体囚徒困境

在中篇中，我们介绍了网络上的两个体囚徒困境博弈，即网络上的个体会与其周围所有的邻居一一进行博弈。更一般地，囚徒困境博弈的参与者可以多于两个，一个典型的模型是  $n$  个体公共品博弈。在此博弈中，每个个体有两种选择——付出一定的代价  $c$  产生  $b$  的收益（即合作，这里  $b/c < n$ ）或者不付出代价也不产生好处（即背叛）。博弈完成后，所有的收益会累加起来并平均分配给所有的参与者，不论他们的选择是什么。此时，在有  $i$  个合作者的集体中，背叛者和合作者的收益分别为

$$P_D(i) = \frac{bi}{n}, P_C(i) = P_D(i) - c.$$

从以上表达式可以看出，对于每个个体来说，选择背叛比选择合作的收益高。在这种情况下，如果每个个体只考虑自身利益最大化，那么所有个体都会选择背叛，以致所有参与者的收益都为 0，这个结果要比每个个体都合作差很多。因此，与篇首提到的两个体捐赠者博弈一样， $n$  个体公共品博弈也能很好地反映个体利益和集体利益之间的冲突。

在有网络结构的群体中，每个个体都会组织一次公共品博弈，参与者是它们自己以及它们的所有邻居。例如，在图 3 中， $C_{11}$  个体总共会参与 4 次博弈，分别由  $C_1, C_{111}, D_{112}$  以及它自己组织；其中  $C_{11}$  个体组织的博弈的参与者有  $C_{11}, C_1, C_{111}$  和  $D_{112}$ ； $C_1$  个体组织的博弈参与者为  $C_1, C_{11}, D_{12}$  和  $D_{13}$ ；以此类推。

在以上的基本模型下，很多学者对网络上的多个体博弈进行了研究，具体可以参考综述 [3]。这里，我们主要介绍最基本的理论分析结果。

对网络上的多个体博弈，一个最基本的问题是在什么条件下，合作者能被自然选择所青睐。

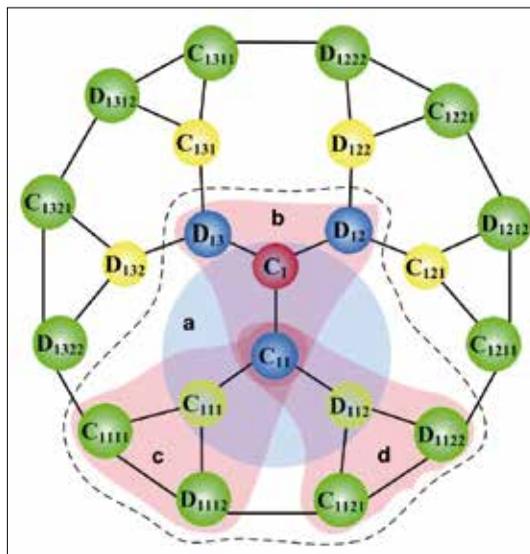


图3 网络上公共品博弈的示意图。源自Li A, Wu B, Wang L. Cooperation with both synergistic and local interactions can be worse than each alone [J]. Scientific Reports, 2014, 4: 5536.

2014年，Li等人的研究回答了此问题：对于大小为  $N$ ，度为  $k$  的随机规则网络上的公共品博弈，当  $N \gg k$  且更新规则为灭生规则时，如果收益代价比  $b/c > (k+1)^2 / (k+3)$ ，那么合作就会被自然选择青睐<sup>[4]</sup>（此结果之后被证明也适用于群体中个体数量无限大的情形<sup>[5]</sup>）。由于群体规模很大的随机规则网络上几乎不存在三元环，也就是个体的两个邻居之间也互为邻居的情况，所以对于拥有聚集系数（clustering coefficient）的网络，还需要进一步的探索。此后，Su等人于2019年拓展了之前的结果，他们得到了在任意聚集系数为  $C$  的传递图（transitive graphs）上，当  $b/c > (k+1)^2 (N-2) / \{N[k+3+(k-1)C]-2(k+1)^2\}$  时，合作行为能被促进<sup>[6]</sup>。

#### 四、基于自我评估的更新 / 决策规则

在以上所介绍的研究中，一个普遍的假设是当个体更新自己的策略时，它会选择以比较收益的方式来模仿其周围的邻居，收益高的个体的策略更有可能被模仿。这种基于模仿的更新规则在自然界是比较普遍的，但这种规则并不是唯一的。近 10 年来的研究表明，当演化模型采用基于模仿的更新规则时，演化结果对模型的细节会很敏感。例如，略微改变一下模仿实施的细节就有可能产生定性上不同的结果。这其中比较著名的例子是 2004 年 Hauert 和 Doebeli 的论文<sup>[7]</sup>（简称 HD 论文）与 1992 年 Nowak 和 May 的论文<sup>[1]</sup>（简称 NM 论文）的对比。具体的对比由 Roca 等人在 2009 年给出：在 NM 论文中，作者发现当更新规则为无条件模仿（即采取自己和周围邻居中收益最高的个体的策略）时，网络结构的存在会促进囚徒困境中合作的演化；但如果 NM 论文使用 HD 论文中所使用的基于收益差的模仿规则时，网络结构的存在反而会抑制合作的演化。同样的结论也可以从 Ohtsuki 等人在 2006 年一系列的文章<sup>[8][9]</sup>中得出：当使用天生规则时，合作有可能被促进；而使用对比较规则（即个体会随机选择一个邻居，然后根据邻居收益与自己收益之差来确定一个概率，并按照该概率模仿邻居的策略）时，合作永远不会被促进。以上的研究结果表明，基于模仿的更新规则导致的演化结果对模型的细节很敏感，也就是说在基于模仿的更新规则下很难得到一致性的结论。当然，这种敏感性在没有网络结构的混合均匀种群也是存在的<sup>[10][11]</sup>。

与此同时，另一种常见但研究甚少的策略更新规则是基于自我评估的规则。在该规则下，个体会比较博弈所得收益与自己内心的期望值，如果收益高于期望值，个体会倾向于维持自己当前的策略，否则个体会切换到其他策略<sup>[12]</sup>。从该规

则的描述中可以看出，在这种规则下，个体在更新策略时受周围邻居策略的影响较小（但影响还是存在的，体现在博弈的过程中）。因此，可以预期在该更新规则下，演化结果受到模型细节的影响将会减小。确实，Du 等人 2014 年的工作<sup>[13]</sup>，Wu 和 Zhou 在 2018 年的工作<sup>[14]</sup>以及 Zhou 等人 2018 年的工作<sup>[15][16]</sup>显示，当个体使用基于自我评估的更新规则时，演化结果对网络结构的变化具有鲁棒性；此外，个体实施更新规则的细节的不同也不会影响最终的演化结果，比如个体期望值的差异<sup>[14][16]</sup>、个体实施自我评估时所用函数的差异<sup>[15][16]</sup>。这与基于模仿的更新规则导致的演化结果形成了鲜明的对比。由此可以看出，基于自我评估的更新规则导致的演化结果对模型的细节具有较强的鲁棒性，得出的结论因此也更具有一致性。

#### 五、动态网络

网络互惠另一个方向的拓展是网络结构与策略的共演化（coevolution），即在演化的过程中，除了个体的策略会随着时间改变之外，网络结构也会随着时间而变化。实际上，相对于静态不变的网络结构，这种拓展是更符合现实的，比如个体与个体之间的社交网络是会随着时间而改变的：当两个个体之间的联系日渐减少时，他们之间的连接可能就消失了；同时，个体也有可能在此过程中与原本不认识的个体建立新的连接。从网络结构的角度来说，这些改变体现在网络上边的变化，简称断边重连。网络上的断边重连可能是随机发生的，也可能是个体有目的地改变自己周围社会环境的一种手段。基于此，这里介绍三种类型的断边重连，分别是基于策略的、基于声望的以及迁移。

基于策略的断边重连指的是不同类型的边（包括 C-C 边，C-D 边和 D-D 边）发生断边重连的概

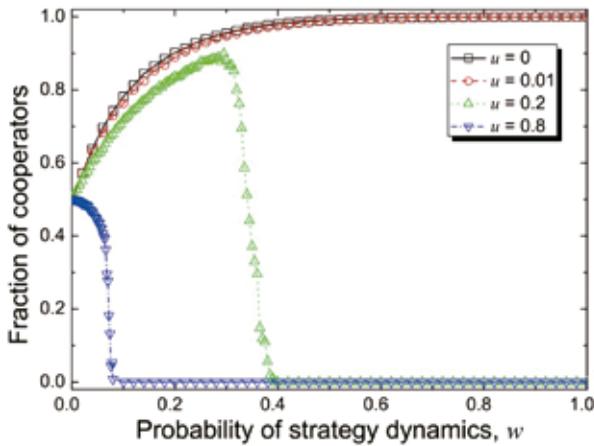


图4 策略更新相对于网络断边重连的相对时间尺度对合作演化的影响。

源自Fu F, Wu T, Wang L. Partner switching stabilizes cooperation in coevolutionary prisoner's dilemma [J]. Physical Review E, 2009, 79(3): 036101.

率可能是不同的。比如，Fu 等人于 2009 年对基于策略的断边重连进行了仿真研究<sup>[17]</sup>。在他们的基本模型中，每一个时间步会随机选择一条 C-D 边，该边连接的两个个体以  $w$  的概率进行策略更新，该边以  $1-w$  的概率进行断边重连。在断边重连的过程中，他们假设只有合作者可以单方面切断与背叛者的连接，并与群体中其他任意一个随机的个体（该个体不能是合作者的已有邻居）建立新的连接。这个基本模型的合理性在于，在合作者与背叛者的交互中，合作者是受剥削的一方，背叛者是得利的一方，此时合作者会有较强的意愿断开连接以免自己被继续剥削。在此模型下，他们发现个体策略更新相对于网络断边重连的时间尺度（即概率  $w$  的大小，当  $w$  接近于 0 时，网络断边重连的速度远快于策略更新的速度）对合作的演化有很重要的影响。具体来说，当代价收益比  $u=c/b$  较大时，个体策略更新相对于网络断边重连的速度越慢越好，也就是  $w$  越小对合作越有利（见图 4）。这是因为在这种情况下，合作者付出的代价较高，如果此时策略更新过快，那么合作者很快就会被背叛者取代；而断边重连的速度很快时，合作者有足够的时间来遴选自己的交互对象，形成合作者团簇，互帮互助，此时就不容

易被背叛者入侵了。

此外，当网络结构的演化速度远快于策略更新的速度时，Wu 等人于 2010 年对该情形的演化动力学进行了理论分析<sup>[18]</sup>。他们的模型每一时间步以  $w$  的概率进行个体的策略更新，以  $1-w$  的概率进行边的断边重连，其中 X-Y 边以概率  $k_{XY}$  发生断边，断边之后，随机选择该边连接的两个个体其中的一个，该个体会随机地与群体中另外一个非邻居个体建立新的连接。在该模型下，他们的理论分析预测当  $w \rightarrow 0^+$  且  $k_{CC} > k_{CD}/(1+u)$  时，合作水平会很低。该理论预测被他们的计算机仿真结果验证，并且他们发现当  $k_{CC}$  越小且  $k_{CD}$  越大时，合作水平越高（见图 5）。该方法对于处理一般的传染病传播过程、观点动力学等动态网络上的传播动力学具有普适性<sup>[19][20]</sup>。

还有一种类型是基于声望的断边重连。它在人类社会中较为普遍，其基本思想是个体可以根据声望来选择自己的交互对象，断开与邻居中声望较低的个体的连接，并与群体中声望较高的个体建立新的连接。Fu 等人于 2008 年对该类型的断边重连进行了研究<sup>[21]</sup>。他们发现的结果与基于策略的断边重连下的结果类似：个体策略更新与基于声望的断边重连的相对时间尺度对结果有很

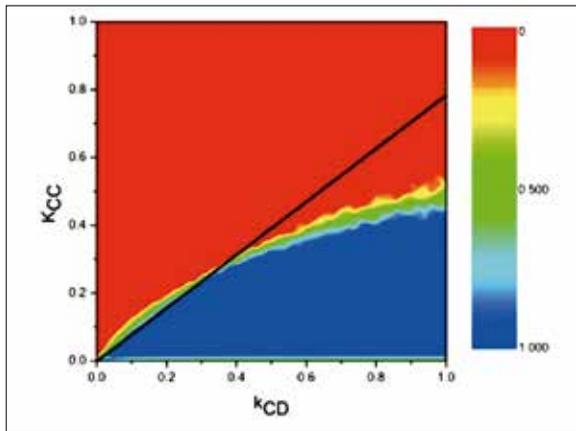


图5 合作水平（深红色代表0，深蓝色代表1）随着断边概率  $k_{CC}$ ,  $k_{CD}$  的变化。黑色实线代表  $k_{CC} = k_{CD} / (1 + u)$ 。  
源自Wu B, Zhou D, Fu F, et al. Evolution of cooperation on stochastic dynamical networks [J]. PLoS One, 2010, 5(6): e11187.

大的影响；并且，当代价收益比（即  $c/b$ ）较大时，较慢的个体策略更新速度更有利于合作的演化。

最后一种类型的断边重连是迁移。在介绍的三种类型中，迁移在自然界中最为普遍，因为它出现在各个层次的生物群体中，从微生物、昆虫、动物到人类。具体来说，当个体不满意当下的环境时，个体就可以选择迁移到其他的环境中。与之前网络上的博弈模型不同的是，此时网络的节点代表的是栖息地，栖息地可以为空也可以被某个个体占据，节点之间的边表示这两个栖息地相邻且处于这两个栖息地的个体可以进行交互。对于该类型的断边重连，Wu 等人于 2011 年在方格网络上对其进行了仿真研究<sup>[22]</sup>。在他们的模型中，除了迁移的基本模型外，他们假设个体会平均以  $s/4$  的概率与它周围的所有邻居进行博弈，博弈完成之后进行策略的更新。他们发现，当引入局部迁移（即个体只能迁移到离自己当前位置较近的空位）时，策略更新相对于博弈的速度越慢（即  $s$  越大），合作水平越高。

当然，以上介绍的三种类型的断边重连只是动态网络研究中的一小部分，其他的类型也有学者进行了研究，比如基于记忆和声望的迁移<sup>[23][24]</sup>、基于推荐的网络结构更新<sup>[25]</sup>等等。关于网络结构

与策略共演化的研究目前已经有很多了，详细的介绍可以参考综述<sup>[26]</sup>。需要注意的是，上述动态网络的演化是基于博弈的，也就是个体的策略或者个体间博弈的结果会对网络的动态变化影响很大。而在现实生活中，个体间交互网络的动态变化可能更多地是由时间和地理位置决定的。对于此类动态网络上合作演化的研究，可参见时序网络（temporal networks）上的相关结果<sup>[27]</sup>，这是一个新的方向。

## 六、小结

在本篇中，我们着重介绍了网络互惠的几个典型拓展，包括交互网络和学习网络的分离，网络上的多个体博弈，基于自我评估的更新规则以及动态网络。这些拓展使得网络互惠的研究更为全面也更为深入。特别地，动态网络的引入使得网络上的演化博弈模型更为接近现实。然而，关于动态网络的绝大多数研究目前采用的都是计算机仿真，相关的理论分析很少（文献 [18][27] 除外），这阻碍了我们更加深入地了解动态网络促进合作演化的原理。关于动态网络的理论建模和分析可以作为下一步网络合作演化动力学的研究方向。

## 参考文献

- [1] Nowak M A, May R. Evolutionary games and spatial chaos [J]. *Nature*, 1992, 359(6398): 826–829.
- [2] Ohtsuki H, Pacheco J M, Nowak M A. Evolutionary graph theory: Breaking the symmetry between interaction and replacement [J]. *Journal of Theoretical Biology*, 2007, 246(4): 681–694.
- [3] Perc M, Gómez-Gardeñes J, Szolnoki A, et al. Evolutionary dynamics of group interactions on structured populations: a review [J]. *Journal of the Royal Society Interface*, 2013, 10(80): 20120997.
- [4] Li A, Wu B, Wang L. Cooperation with both synergistic and local interactions can be worse than each alone [J]. *Scientific Reports*, 2014, 4: 5536.
- [5] Li A, Wang L. Evolutionary dynamics of synergistic and discounted group interactions in structured populations [J]. *Journal of Theoretical Biology*, 2015, 377: 57–65.
- [6] Su Q, Li A, Wang L, et al. Spatial reciprocity in the evolution of cooperation [J]. *Proceedings of the Royal Society B*, 2019, 286(1900): 20190041.
- [7] Hauert C, Doebeli M. Spatial structure often inhibits the evolution of cooperation in the snowdrift game [J]. *Nature*, 2004, 428(6983): 643–646.
- [8] Ohtsuki H, Hauert C, Lieberman E, et al. A simple rule for the evolution of cooperation on graphs and social networks [J]. *Nature*, 2006, 441(7092): 502–505.
- [9] Ohtsuki H, Nowak M A. The replicator equation on graphs [J]. *Journal of Theoretical Biology*, 2006, 243(1): 86–97.
- [10] Wu B, García J, Hauert C, et al. Extrapolating weak selection in evolutionary games [J]. *PLoS Computational Biology*, 2013, 9(12): e1003381.
- [11] Wu B, Bauer B, Galla T, et al. Fitness-based models and pairwise comparison models of evolutionary games are typically different—even in unstructured populations [J]. *New Journal of Physics*, 2015, 17(2): 023043.
- [12] Du J, Wu B, Altrock P M, et al. Aspiration dynamics of multi-player games in finite populations [J]. *Journal of the Royal Society Interface*, 2014, 11(94): 20140077.
- [13] Du J, Wu B, Wang L. Aspiration dynamics in structured population acts as if in a well-mixed one [J]. *Scientific Reports*, 2015, 5: 8014.
- [14] Wu B, Zhou L. Individualised aspiration dynamics: Calculation by proofs [J]. *PLoS Computational Biology*, 2018, 14(9): e1006035.
- [15] Zhou L, Wu B, Vasconcelos V V, et al. Simple property of heterogeneous aspiration dynamics: Beyond weak selection [J]. *Physical Review E*, 2018, 98(6): 062124.
- [16] Zhou L, Wu B, Du J, et al. Aspiration dynamics generate robust predictions in structured populations. arXiv preprint arXiv:1810.00387. 2018.
- [17] Fu F, Wu T, Wang L. Partner switching stabilizes cooperation in coevolutionary prisoner's dilemma [J]. *Physical Review E*, 2009, 79(3): 036101.
- [18] Wu B, Zhou D, Fu F, et al. Evolution of cooperation on stochastic dynamical networks [J]. *PLoS One*, 2010, 5(6): e11187.
- [19] 王龙, 武斌, 杜金铭, 等. 复杂动态网络上的传播行为分析——献给清华大学郑大钟教授. *中国科学: 信息科学*, 2020, 50.
- [20] Wu B, Du J, Wang L. Bridging the gap between opinion dynamics and evolutionary game theory: some equivalence results [C]. 2020 39th Chinese Control Conference (CCC). 2020: 6707–6714.
- [21] Fu F, Hauert C, Nowak M A, et al. Reputation-based partner choice promotes cooperation in social networks [J]. *Physical Review E*, 2008, 78(2): 026117.
- [22] Wu T, Fu F, Wang L. Moving away from nasty encounters enhances cooperation in ecological prisoner's dilemma game [J]. *PLoS One*, 2011, 6(11): e27669.
- [23] Li A, Wu T, Cong R, et al. One step memory of group reputation is optimal to promote cooperation in public goods games [J]. *EPL (Europhysics Letters)*, 2013, 103(3): 30007.
- [24] Cong R, Wu B, Qiu Y, et al. Evolution of cooperation driven by reputation-based migration [J]. *PLoS One*, 2012, 7(5): e35776.
- [25] Wu B, Park H J, Wu L, et al. Evolution of cooperation driven by self-recommendation [J]. *Physical Review E*, 2019, 100(4): 042303.
- [26] Perc M, Szolnoki A. Coevolutionary games—a mini review [J]. *BioSystems*, 2010, 99(2): 109–125.
- [27] Li A, Zhou L, Su Q, et al. Evolution of cooperation on temporal networks [J]. *Nature communications*, 2020, 11(1): 1–9.